# Mechanisms of saccadic decision making while encoding naturalistic scenes

# Mechanisms of saccadic decision making while encoding naturalistic scenes

**R. Calen Walshe**

Psychology Department, University of Edinburgh, Edinburgh, UK

**Antje Nuthmann**

Psychology Department, University of Edinburgh, Edinburgh, UK

**Saccadic eye movements are the primary vehicle by which human gaze is brought in alignment with vital visual information present in naturalistic scenes. Although numerous studies using the double-step paradigm have demonstrated that saccade preparation is subject to modification under certain conditions, this has yet to be studied directly within a naturalistic scene-viewing context. To reveal characteristic properties of saccade programming during naturalistic scene viewing, we contrasted behavior across three conditions. In the Static condition of the main experiment, double-step targets were presented following a period of stable fixation on a central cross. In a Scene condition, targets were presented while participants actively explored a naturalistic scene. During a Noise condition, targets were presented during active exploration of a 1/*f* noise-filtered scene. In Experiment 2, we measure saccadic responses in three Static conditions (Uniform, Scene, and Noise) in which the backgrounds are the same as Experiment 1 but scene exploration is no longer permitted. We find that the mechanisms underlying saccade modification generalize to both dynamic conditions. However, we show that a property of saccade programming known as the saccadic dead time (SDT), the interval prior to saccade onset during which a saccade may not be canceled or modified, is lower in the Static task than it is in the dynamic tasks. We also find a trend toward longer SDT in the Scene as compared with Noise conditions. We discuss the implication of these results for computational models of scene viewing, reading, and visual search tasks.**

## Introduction

To access information-rich regions of the visual field, the visual system engages in eye-movement behaviors known as fixations and saccades. The coordination of such movements involves a complex array of motor control mechanisms operating on distinct spatio-temporal scales. Fixations are defined as the state at which the eye remains in a relatively stable position on some aspect of the visual environment. In scene perception, it is known that the duration of fixations are dependent on such factors as the type of task that people are engaging in (Mills, Hollingworth, Van der Stigchel, Hoffman, & Dodd, 2011; Nuthmann, Smith, Engbert, & Henderson, 2010; Smith & Mital, 2013), the relevance of the fixated material to the task goals (Land & Hayhoe, 2001), and the lower-level stimulus properties such as the luminance (Henderson, Nuthmann, & Luke, 2013; Walshe & Nuthmann, 2014) and color (Ho-Phuoc, Guyader, Landragin, & Guérin-Dugué, 2012) of the scene. Saccadic eye movements are the primary mechanism by which the eye is brought into spatial alignment with scene content that is to be inspected in high-resolution foveal vision. Where observers fixate is influenced by both mid-level and higher-level stimulus properties (Nuthmann & Einhäuser, 2015). However, it is also known that the effect of image features on fixation selection in scenes may be overridden by task demands (Einhäuser, Rutishauser, & Koch, 2008). Like fixation durations, properties of saccades are also known to vary depending on the task. For instance, participants generate larger saccade amplitudes when searching for an item than when memorizing a scene for later recall (Mills et al., 2011).

Theories of eye-movement control are primarily interested in elucidating the underlying, hidden mechanisms that generate behaviors such as saccades and fixations. The question of what event during stimulus processing results in the initiation of a saccade program to shift fixation away from the currently fixated location is a matter of considerable debate. This question has been most directly addressed in theories of eye-movement control while engaging in reading

Citation: Walshe, R. C., & Nuthmann, A. (2015). Mechanisms of saccadic decision making while encoding naturalistic scenes. *Journal of Vision*, *15*(5):21, 1–19, http://www.journalofvision.org/content/15/5/21, doi:10.1167/15.5.21.

doi: 10.1167/15.5.21 Received July 31, 2014; published April 29, 2015 ISSN 1534-7362 

behavior in which two competing views have been suggested (Reingold, Reichle, Glaholt, & Sheridan, 2012, for review). The first view suggests that a saccade program is triggered only once the currently viewed stimulus has been processed to a sufficient degree (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Pollatsek, & Rayner, 2012). Variants of this view are referred to as *cognitive trigger theories*, as saccadic eye movements are generated only when cognition-related processing has been achieved to a sufficient depth. In contrast to the triggering mechanisms just described are those that suggest that the variability in the termination of a fixation is a result of difficulties in lexical processing that interfere with saccade initiation processes. In such models, the decision to initiate a saccade is achieved by an autonomous random timer, and the duration of this timing process may be modulated by the difficulties encountered during stimulus processing (Engbert, Nuthmann, Richter, & Kliegl, 2005; Nuthmann et al., 2010; Trukenbrod & Engbert, 2014). Therefore, moment-to-moment difficulties in processing result in increased random timing intervals and, consequently, longer fixation durations. Such principles of saccade generation have been adapted to explain fixation duration distributions within complex scene-viewing tasks. The CRISP model (Nuthmann et al., 2010) incorporates a random-walk timer that generates signals to begin the programming of a saccade. Importantly, the CRISP model also allows such saccade programs to be canceled in the case that programming has not proceeded to a sufficiently advanced stage. The CRISP model has been demonstrated to provide good fits to empirical data under a number of experimental conditions (Nuthmann et al., 2010; Nuthmann & Henderson, 2012).

A powerful tool that has been used to reveal the empirical properties of saccade programming timelines is known as the double-step paradigm (Westheimer, 1954; Becker & Jürgens, 1979). Classic variations of the double-step paradigm involve presenting participants with two targets along a horizontal axis with a varying interstimulus interval separating the two targets. For instance, in a seminal study, Becker and Jürgens (1979) had a condition in which a first target was presented at 15° to the left or right of fixation with a second target presented at 30° in the same direction at delays of 50, 100, 150, and 200 ms. The participants' task was to fixate the target as quickly as possible, which thereby placed a saccade program that had been initiated to target the first location in competition with a saccade program targeting the second location. The critical behavioral measure in this study was the amplitude of the saccade that was observed in response to the presentation of the double-step stimulus. Responses observed to fall spatially close to the location of the initial target step permitted the inference that information regarding the updated target location was not incorporated into the programming of the saccadic response. Likewise, the degree to which saccade amplitudes deviated from the initial location and tended to land close to the second target provided a measurement of the degree to which the second target had influenced the resulting saccade. The authors observed an interesting temporal dependency between the amplitude of the saccadic response and the presentation of the second target. If the response saccade occurred in close temporal proximity to the appearance of the second target, then the response saccade tended to fall close to the initial target position. As the temporal interval between the presentation of the second target and the response increased, responses gradually tended toward the final target position. These results revealed that saccades may be reprogrammed when updated position information is available and that the mechanisms underlying this behavior are dependent on the temporal relationship between the updated target position and the onset of the response saccade. This finding has since been replicated by many authors (Aslin & Shea, 1987; Camalier et al., 2007; Findlay & Harris, 1984; Ludwig, Mildinhall, & Gilchrist, 2007; Ray, Schall & Murthy, 2004).

A question that follows directly from such results is at what point in time prior to the onset of a saccade does updated target position information cease to have an influence on the impending saccade? Such a point of no return is often referred to in the literature as saccadic dead time (SDT). Findlay and Harris (1984) analyzed double-step data from a replication of Becker and Jürgens (1979) and found that saccades begin to incorporate targets that are displaced 80 ms prior to saccade onset. More recently, Ludwig et al. (2007) used a double-step task to investigate whether estimated values for SDT show evidence for sensitivity to task effects. In their study, the angle of displacement between the first and second double-step target was manipulated. They found that SDT increased as the angle between the initial and final target locations increased. These results provide empirical support for the claim that SDT values are sensitive to characteristics of the task environment.

The concept of a point of no return has been suggested as an explanation for a number of empirical effects observed within the scene-viewing literature. In a procedure known as the stimulus onset delay (SOD) paradigm (Henderson & Pierce, 2008; Shioiri, 1993) participants viewed naturalistic scene stimuli while engaged in an encoding task. During a critical fixation, the scene was removed and replaced by a mask, thereby removing scene content from further processing. Within the same fixation, the scene was then restored to view at varying delays. Results from the SOD paradigm showed a bimodal distribution of fixation durations

such that one population was independent of the length of the delay whereas the other population increased in proportion to the length of the delay. In the context of the CRISP model, Nuthmann et al. (2010) suggested that such bimodality arises due to two factors. First, when the scene disappears from view, the rate at which saccade programs are generated slows down to reflect the lack of incoming visual information. Second, if a saccade program is currently within the labile stage of saccade programming (i.e., has not passed the point of no return), then the current saccade program is subject to stochastic cancellation. Together, these processes combine to yield delays in the programming of saccades and therefore longer fixation durations. Wu et al. (2013) also demonstrated an influence of a point of no return on saccade programming in a study in which a scene was swapped with a novel scene while participants were engaged in a viewing task. In the saccade that immediately followed the change, it was found that participants systematically programmed saccades to target the center of the screen. However, such a center scene bias occurred primarily for saccades that were initiated at relatively long durations after the change. For saccades that were initiated shortly after the display change, there was no evidence for such a bias. Similar to the SOD paradigm, such a result is well accounted for by the fact that in those saccades that did not target the center, the scene change occurred when saccade programming had already passed the point of no return and could therefore no longer influence saccade programming.

The principles derived from such investigations with double-step stimuli have provided the basis for the implementation of eye-movement control models in scene viewing as well as in reading. In the E-Z Reader model (Reichle et al., 1998), the authors introduced the concepts of a labile and a nonlabile stage of saccade programming. This distinction implies a multistage saccade programming architecture whereby saccade programming that is within the labile stage is subject to cancellation. However, once programming has progressed beyond the labile stage into the nonlabile stage, cancellation is no longer possible. Such multistage saccade programming assumptions have since been incorporated into a number of models that explain oculomotor control under a variety of task conditions such as scene viewing (Nuthmann et al., 2010), reading (Engbert et al., 2005; Pollatsek, Reichle, & Rayner, 2006; Reichle et al., 1998), and visual search (Trukenbrod & Engbert, 2014).

One notable difference between these models is the assumptions that are made regarding the duration of the nonlabile stage. In both the CRISP model (Nuthmann et al., 2010) and the E-Z Reader model (Reichle et al., 1998), it is assumed that the nonlabile stage adopts a fixed duration and is not dependent on stimulus characteristics. On the other hand, as of SWIFT-II (Engbert et al., 2005), the model makes the explicit assumption that the duration of the nonlabile stage may vary depending on the amplitude of the planned saccade. Therefore, the SWIFT model explicitly incorporates systematic variability in SDT while the CRISP model and E-Z Reader do not. It is important to note that all models just introduced do incorporate *random* (i.e., unsystematic) variability in the duration of the nonlabile stage as the duration of a given nonlabile stage is drawn from a statistical distribution. Although Ludwig et al. (2007) provides some empirical support for stimulus-dependent differences in the duration of the nonlabile stage, there is currently no empirical research demonstrating task-specific differences within reading or scene-viewing contexts. Furthermore, upon inspection of the model parameters that are used to describe the duration of the nonlabile stage, it becomes evident that there is very little consistency between (and within) models regarding the duration of this stage. We return to this issue in the Discussion.

The aim of the current study is to investigate saccade programming during naturalistic scene perception by embedding a double-step task within scene-viewing contexts. Two primary questions are addressed with these experiments. The first addresses an empirical gap in the scene-viewing literature by testing whether saccade cancellation operates in an analogous manner within scene viewing as it does within classic double-step investigations. The assumption that findings from low-level tasks generalize to high-level task contexts has often been made by models of oculomotor control (Engbert et al., 2005; Nuthmann et al., 2010; Reichle et al., 1998; Trukenbrod & Engbert, 2014). However, it is an open question that has remained largely unaddressed (but see Walshe & Nuthmann, 2013). Our second question asks whether SDT values vary depending on experimental context. By deriving such an estimate, we also explore how such a value may provide an indirect measure of the duration of the nonlabile stage by taking into account the delays in transmission of information between the retina and cortical regions responsible for saccadic decisions.

In Experiment 1, we address these questions by comparing double-step performance across three conditions. The first condition (Static) replicates a classic version of the double-step procedure (Becker & Jürgens, 1979). In a second condition (Scene), a scene-viewing double-step task is conducted by presenting the double-step targets during active exploration of scene content. The third condition (Noise) replicates the experimental design of the scene-viewing double-step task but instead replaces the scene with a phase noise–transformed background stimulus. Such a transformation is achieved by applying noise to the phase spectrum but leaves the amplitude spectrum intact

(Einhäuser et al., 2006). Applying a transformation in this manner removes object and other higher-order scene statistics yet retains the $1/f$ characteristics of the amplitude spectra that is typical of naturalistic scenes (Einhäuser et al., 2006). In a follow-up experiment, we test how SDT is influenced by scene background independently of dynamic movement context. In Experiment 1, observed differences in SDT between the Static condition and the dynamic movement conditions (Scene, Noise) could be attributed to either the background content or differences in the movement context. Experiment 2 provides a test of these two hypotheses by comparing the three backgrounds used in Experiment 1 in static movement contexts only.

To make comparative inferences about saccade programming across conditions, we use a nonlinear mixed-effects (*nlme*) regression framework that improves on previous methods of analyzing double-step performance. Population-level parameters provided a method to compare performance across conditions, whereas individual parameter estimates were extracted from the fitted model and were used to provide a by-participant measure of SDT in the three conditions.

# Experiment 1

## Methods

### Participants

Two men and 13 women (mean age = 23 years) recruited from the University of Edinburgh student population participated in the study. Participants completed all experimental conditions in one session, which lasted approximately 1.5 h. Each participant was paid £7 per hour of participation in compensation for their time. The study conformed to the tenets of the Declaration of Helsinki, and written consent was supplied by the participants prior to the experiment.

### Apparatus

Stimuli were presented on a 21-inch CRT monitor with a refresh rate of 140 Hz, and the monitor screen was at a distance of 67 cm from the participant. During stimulus presentation, participants' eye movements were recorded using an SR Research EyeLink 1000 Desktop mount system. It was equipped with the 2000-Hz camera upgrade, allowing for binocular recordings at a sampling rate of 1000 Hz for each eye with an average spatial accuracy of 0.25°–0.5° of visual angle. Viewing was binocular, and both eyes were tracked. Only the position of the right eye was used in the analysis. A chin rest was used to achieve stability of a participants' head position relative to the screen. The experiment was implemented in MATLAB 2009b using the OpenGL-based Psychophysics Toolbox 3 (Brainard, 1997; Kleiner, Brainard, Pelli, Ingling, Murray, & Broussard, 2007), which incorporates the EyeLink Toolbox extensions (Cornelissen, Peters, & Palmer, 2002). The software allowed precise control over the timing of display changes. To detect fixations online, we implemented a nine-sample online velocity estimation algorithm in MATLAB that aimed to mimic Data Viewer's offline velocity estimation procedure (SR Research Ltd., 2006). Fixations were detected offline using SR Research Data Viewer to parse the gaze samples into sequences of fixations and saccades.

### Stimuli

In all conditions, the targets consisted of isoluminant 1.5° square boxes presented in the color pink ($CIE_{Lab}$ L = 65.48, a = 61.84, b = −26.03). In the Static task, the background was uniformly black. When participants were required to fixate on a central cross, it was presented in red ($CIE_{Lab}$ L = 53.23, a = 80.42, b = 66.96). In the Scene task, participants viewed images of 200 naturalistic scenes, in addition to four practice scenes. Each scene had a resolution of 800 × 600 pixels and was presented in full color. Scenes were collected from online databases such as Flickr and Google images. They were selected to include a variety of categories such as indoor and outdoor as well as urban and nature scenes. Each scene was viewed by the participant only once over the duration of the experiment. At a viewing distance of 67 cm, the scenes subtended 33° × 25°. The stimuli in the Noise task consisted of 200 background images that were constructed by applying a phase-noise filtering procedure to the images used in the Scene task. The application of the phase noise–filtering process allowed removal of higher-order scene statistics such as those used to determine edges and contours while leaving the amplitude spectrum unmodified. Phase noise stimuli were created by transforming the original scene images into Fourier space, where additive noise drawn from a uniform distribution was added to the phase spectrum (Einhäuser et al., 2006). An inverse Fourier transformation was then applied to the images to convert them back to image space.

### Procedure

Three double-step tasks were conducted to compare characteristics of saccade programming in static and dynamic gaze conditions. The order of the conditions was counterbalanced across participants. In each of the three tasks, participants completed 200 trials that were classified as either double-step (40%) or single-step (60%) trials. Single-step trials were included to ensure that

participants could not make accurate anticipatory saccades to the final resting location of the double-step target. On single-step trials, the center of the 1.5° pink box was presented at an eccentricity of 7° from the point at which the participant was currently fixating. The order of single-step and double-step trials was randomized.

On double-step trials, two targets were presented at distances of 7° and 14° of visual angle from the fixated location. The first target is referred to as the *initial target* and the second target as the *final target*. Both targets were always presented in the same direction relative to the fixation location. Therefore, because the initial target was presented at 7° from the fixation location, the final target was always presented at a further 7° in the same direction. The color and luminance of the stimuli in the double-step condition were identical to that of the target in the single-step condition. The timing of the initial target step was identical to that of the timing of the single-step stimuli described above. The initial target step was presented for varying amounts of time prior to the onset of the final target step. The interval defining the amount of time elapsed between the onset of the first target step and the onset of the second target step is referred to as the *target step delay* (TSD). The final target was presented simultaneously with the disappearance of the first target. Therefore, the subjective impression of this procedure is that the first target step *jumps* to the second target location.

The TSD for a trial was defined in an adaptive manner such that TSD varied depending on the amplitude of the response on the previous double-step trial (Camalier et al., 2007). A *compensated saccade* refers to a saccade that was programmed to go directly to the final target location, whereas a *noncompensated saccade* is one in which the saccade was programmed to the initial target location. From previous investigations, it is known that short TSDs tend to result in final target response saccades whereas longer TSDs tend to result in initial target responses (Becker & Jürgens, 1979; Camalier et al., 2007). Following a double-step trial in which a compensated saccade was produced, TSD was increased and TSD was decreased following noncompensated saccades. The adaptive increment was defined as 50 ms; a lower bound on TSD was set at 50 ms. The purpose of this adaptive procedure was to balance the number of compensated and noncompensated response saccades. Compensated saccades were detected online and were identified when a saccade was made within 2° of the final target. The presentation of the stimulus was synced with the vertical retrace of the monitor. Full presentation of the stimulus was therefore delayed from the defined TSD by up to 7.14 ms. The trial was terminated 1,000 ms following the first saccade made in response to the presented targets.

### Static task

In the Static condition, participants initiated a trial by fixating on a cross presented at the center of the screen. If the eye tracker could not reliably detect fixation on the cross, a recalibration routine was initiated. Following a randomly timed delay of 2000 to 3000 ms, targets were placed on the same horizontal axis as the fixation cross and were presented either to the left or to the right. On double-step trials, the final target was always presented on the same side of the cross as the initial target. Targets were presented to the left or right with equal frequency, and the side was randomly selected for a particular trial. Participants were instructed to fixate the box as quickly and accurately as possible. The design in the Static condition is visualized in Figure 1.

### Dynamic tasks

In the dynamic tasks (Scene, Noise), participants initiated a trial by fixating on a central cross. Once the cross had been fixated, the image was revealed and the participant was free to make unconstrained fixations on the screen. In both conditions, the task was to encode the presented image for later recall. Participants were instructed that the recall phase would commence once all the stimuli had been presented. Once the participant had made 10 saccades on the current trial, a critical fixation was identified and the targets were presented. The targets were presented immediately upon detection of the onset of a critical fixation. Fixations were detected online using a custom nine-point velocity estimation algorithm (see the Apparatus section). A primary difference between the static and dynamic tasks is the trajectory at which the targets were presented. In contrast to the Static task, in the dynamic tasks the targets could be presented along any axis and the location of the targets was determined by the position of the current and preceding fixation locations. The targets were presented along an axis that was derived by intersecting the coordinates of the current fixation with the previous fixation. The initial target was therefore placed at a distance of 7° from the currently fixated location in the same direction as the most recent saccade (see Figure 1 for details). The decision to place the targets along such a trajectory was done to control for systematic viewing biases that are known to exist during the exploration of naturalistic scenes. First, it is known that during scene-viewing tasks, saccades are most frequently programmed in the horizontal direction (Foulsham, Kingstone, & Underwood, 2008; Nuthmann & Henderson, 2010; Tatler & Vincent, 2008). We confirmed that such a systematic bias exists in our dynamic conditions by visual inspection of radial histograms showing saccade angles relative to the horizon (see Figure 2). To account for

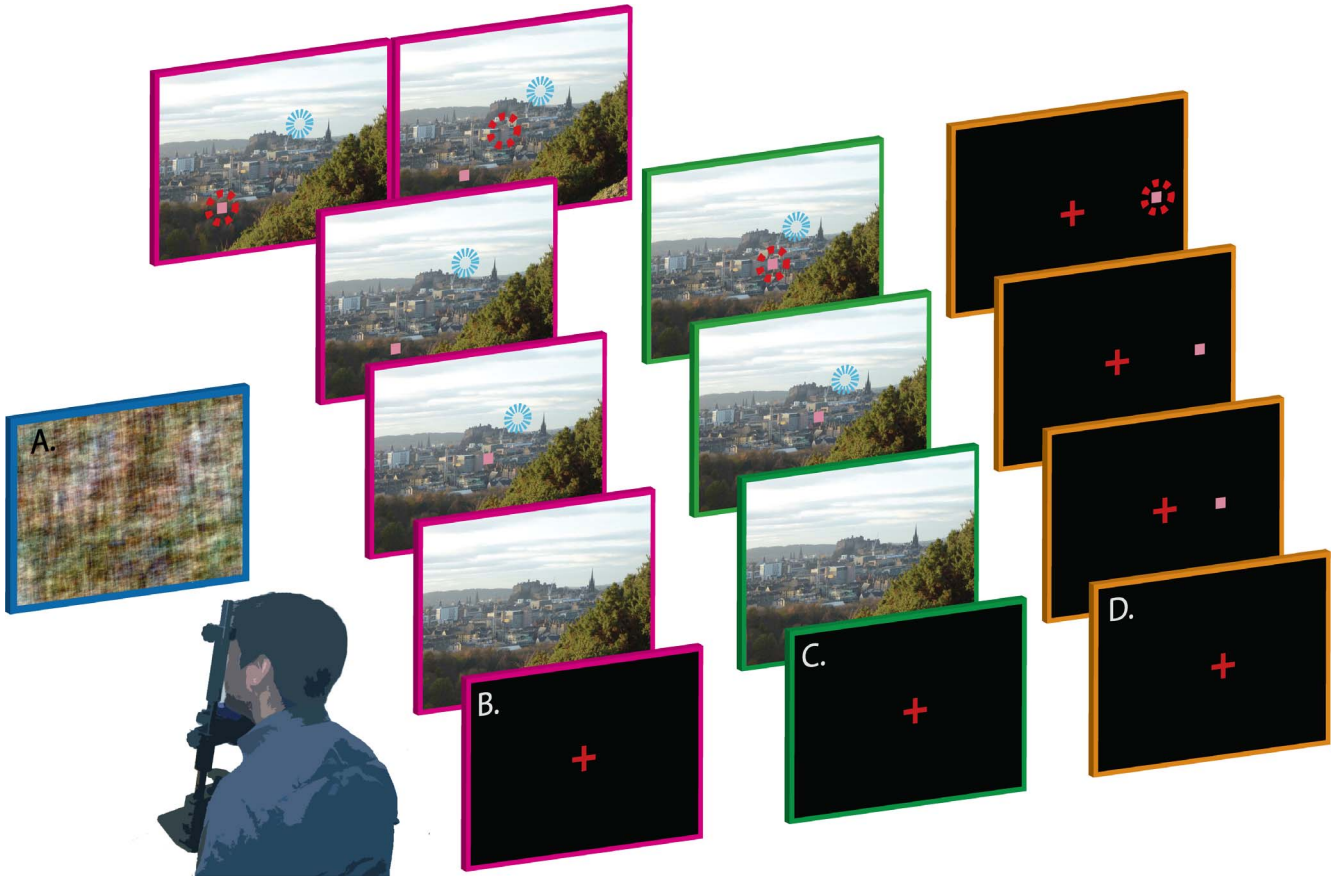


Figure 1. Experiment 1 design. (A) An example stimulus from the 1/*f* Noise condition is presented. The procedure for the double-step condition in the dynamic tasks is illustrated in (B). The experiment begins with the participant fixating on a central cross. The participant then views a scene (or 1/*f*) stimuli for later recall. On the 10th fixation, the first target is presented immediately upon detection of fixation onset, followed by the second target at intervals of 50, 100, 150, or 200 ms. An adaptive staircase method was used to define the interval on a specific trial. The dashed blue circle represents an example fixation location before target presentation, and the red circle represents an example fixation location after a target response is generated. A *compensated saccade* occurs when a saccadic response is generated directly to the final position of the target (top left of panel B). A *noncompensated saccade* occurs when a response is erroneously generated to the first target position (top right of panel B). (C) The single-step condition is illustrated, in which only a single target is presented. (D) The procedure for a double-step trial in the Static condition. Participants fixate on a red cross at the center of the screen. Targets then appear at intervals of 2000 to 3000 ms. Trials can either be single-step or double-step trials, and the timing is the same as in panels B and C.

such a horizontal bias, we implemented a static double-step task in which targets were presented exclusively on the horizon. This was done to maximize the match between target trajectories in the static and dynamic conditions. Second, we wanted to control for the effect of *saccadic momentum*. Saccadic momentum refers to the tendency for fixation durations to increase as the angle between the preceding and subsequent saccade direction increases (Wilming, Harst, Schmidt, & König, 2013). For example, saccades that are programmed perpendicular to the direction of the preceding saccade result in longer fixation durations. Therefore, by placing both targets along the same trajectory, any influence on the timing of the saccade due to saccadic momentum is minimized. The design in the dynamic tasks is visualized in Figure 1.

## Results

Prior to analysis, the data were preprocessed to exclude cases that did not conform to sufficient data quality standards. Each participant viewed 200 trials in each condition, 40% (80) double-step trials and 60% (120) single-step trials. In double-step trials, we excluded any trial in which the response saccade was initiated prior to the time at which both targets had been presented. This left an average of 61 double-step trials in the Static condition, 74 in the Scene condition, and 72 in the Noise condition. Saccades that were clearly too short (<2°) to be considered as responses to either of the targets were excluded. Furthermore, if a blink occurred immediately before or immediately after the fixation in which the targets were presented, this

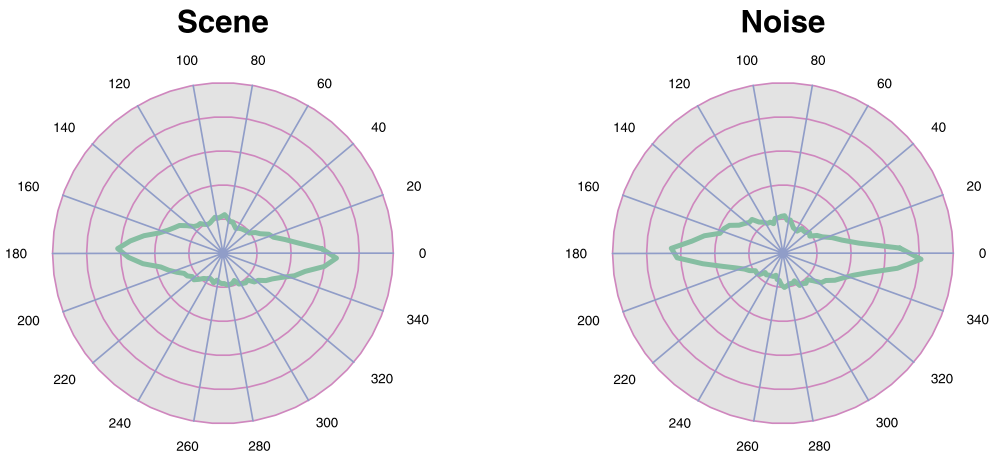


Figure 2. Distribution of saccade direction in Experiment 1. In both the Scene and Noise conditions, there is a clear preference to program saccades along the horizontal axis. Angles of 0° and 180° indicate saccades programmed along the horizon. Densities were calculated from a bin size of 7.2°.

trial was excluded. In the Static task, if the targets were presented when the fixation deviated by more than 2° from the fixation cross, this trial was excluded. An additional criterion was defined such that if a saccade was programmed more than 45° away from the direction that the targets were presented in, these saccades were considered not to be programmed in response to detection of the targets and were excluded from the analysis. After all exclusion criteria had been applied, there remained an average of 56 double-step trials in the Static condition, 67 trials in the Scene condition, and 68 trials in the Noise condition.

## Modeling the amplitude transition function

The amplitude transition function (ATF) relates the resulting saccadic response amplitude to a quantity referred to as delay (D). D measures the time elapsed between the onset of the second target step and the onset of the response saccade (Becker & Jürgens, 1979). The amplitude of the saccade provides information about whether a saccade was successfully reprogrammed or not. In the present experiment, amplitudes of approximately 7° indicate saccades programmed to the initial target, whereas amplitudes greater than 7° can provide evidence for the saccade being executed toward the final target location.

Values of D provide a measurement of how much time has elapsed between the onset of the second target and the onset of the saccadic response. Therefore, D measures the amount of time available to reprogram a saccade to the new target location. If the value of D is high, this implies that the second target was available for a relatively long period prior to the onset of the saccade. When the value of D is low, the saccade was executed very shortly after the presentation of the second target. By combining D with the amplitude of the response, thereby constructing an ATF, it is possible to ask the following question: What is the minimal amount of time prior to the onset of the saccade that the second target must be presented to have an influence on the resulting saccade?

Previous research has demonstrated that the ATF in double-step tasks may be well described by a curve that closely resembles the logistic function (Becker & Jürgens, 1979; Ludwig et al., 2007). This was confirmed by graphical analysis of the data in the present experiments as well as ATFs previously reported in the literature. The following four-parameter logistic function was used to model response amplitude as a function of D in the three experimental conditions:

$$f(x) = \alpha + \frac{\beta - \alpha}{1 + e^{\gamma(\delta - x)}}, \quad (1)$$

where $\alpha$ represents the lower bound for the logistic function, $\beta$ represents the upper bound, $\gamma$ is a scaling parameter, and $\delta$ defines the inflection point.

As a novel approach, ATFs were estimated with a nonlinear mixed-effects regression framework using the *nlme* package (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2014) implemented in the R software for statistical computing (R Development Core Team, 2012). Using this approach, we model the variability in the estimation of parameters contributed by both individual participants and experimental condition. Specifically, a nonlinear mixed-effects model provides a method of simultaneously estimating the fixed (population-level) parameters and the random (individual-level) parameters (Pinheiro & Bates, 2000). A benefit of estimates derived from such a framework is that participant-level parameter estimates are weighted by their corresponding population parameters, providing a measure of protection against overfitting the ATFs to

| Fixed effects | | | | Random effects | |
|---|---|---|---|---|---|
| Parameter | Estimate | SE | *t* value | Parameter | $\sigma$ |
| Lower bound ($\alpha$) | | | | | |
| Static (intercept) | 6.15 | 0.19 | 31.56 | Static | 0.59 |
| Noise | −0.07 | 0.18 | −0.39 | Noise | 0.42 |
| Scene | 0.164 | 0.173 | 0.946 | Scene | 0.20 |
| Upper bound ($\beta$) | | | | | |
| Static (intercept) | 13.05 | 0.15 | 85.62 | Static | 0.47 |
| Noise | −1.18 | 0.21 | −5.50 | Noise | 0.58 |
| Scene | −0.68 | 0.24 | −2.81 | Scene | 0.70 |
| Inflection ($\delta$) | | | | | |
| Static (Intercept) | 90.05 | 2.30 | 39.10 | Static | 6.77 |
| Noise | 16.49 | 2.31 | 7.14 | Noise | 4.15 |
| Scene | 33.10 | 2.10 | 11.04 | Scene | 6.80 |
| Scale ($\gamma$) | | | | | |
| Static (Intercept) | 0.09 | 0.01 | 8.39 | Static | 0.03 |
| Noise | 0.07 | 0.01 | 1.94 | Noise | 0.01 |
| Scene | −0.02 | 0.02 | −1.26 | Scene | 0.05 |
| Error term | | | | $\varepsilon$ | 1.64 |

Table 1. ATF regression model (Experiment 1). *Notes:* Estimated parameters for the four-parameter logistic regression model. Means, standard errors, and *t* values of fixed effects; standard deviations of the random effects.

individual-level data. By-participant random effects (intercept and slope) were included for all four parameters of the logistic function, thereby allowing all parameters of the model to vary in a by-participant manner. The random effect covariance matrix was assumed to follow a block diagonal structure. That is, correlations between random intercepts and slopes were permitted only when grouped within the logistic parameters $\alpha$, $\beta$, $\gamma$, and $\delta$ and were assumed to be 0 otherwise. For example, correlations between the upper-bound intercept and slopes were estimated, but correlations between the upper- and lower-bound random effects were assumed to be 0. Parameters were estimated by maximizing the log-likelihood of the model given the observed responses. The parameters of the fitted model are presented in Table 1.

## Saccade modification

The ATF is a direct way to measure the behavioral correlates of mechanisms underlying saccade reprogramming as it measures the amplitude of saccades that are initiated under conditions in which updated target position information is available. Furthermore, the ATF plots the response amplitude as a function of the duration that has elapsed since the presentation of the second target (D). Therefore, the ATF allows the relationship between the time spent processing the second target and the amplitude of the response to be revealed. The ATF allows specific predictions about saccade cancellation processes to be tested. In the case that saccades can never be modified, the ATF predicts a flat function with an intercept close to the location at which the first target is presented. In the present experiment, the predicted intercept would be 7°. In the case in which modification is always possible, a flat function would also be predicted, but in this case, the intercept would be predicted to be near the location at which the second target is presented or 14° in the present experiment. An intermediate hypothesis between these two extremes is that modification of a saccade program is possible and becomes increasingly likely at greater temporal separation between the onset of the second target and the onset of the saccadic response. In this case, a monotonically increasing ATF would be predicted with a lower asymptote located close to the first target location that gradually increases and asymptotes close to the final target location. Inspection of Figure 3 reveals that the shape of the ATF in all three experimental conditions conforms to the intermediate hypothesis. Saccade programs can be modified, and as greater values of D are observed, a saccade targeting the final location becomes increasingly likely.

Furthermore, the regression modeling reveals that, for double-step trials, saccades significantly undershoot the targets and that the degree of undershoot depends on experimental condition. The Static condition was selected as the intercept for each of the logistic parameters that were estimated. Effects were determined to be significantly different from 0 when $|t| > 1.96$ was observed. For the lower bound, the estimated amplitudes for all conditions were less than the distance at which the targets were placed (7°). Task did not significantly influence the lower bound. That is, the estimated change in lower bound from the Static condition (intercept) was not significant for either the Noise or Scene condition. Significant task effects were observed for the estimated upper bound. The upper bound estimate for the Static condition (intercept) was $\hat{\beta} = 13.05°$, $t = 85.62$, which indicates an undershoot to the second target. The decrease in upper bound was greatest in the Noise condition, $\hat{\beta} = -1.18°$, $t = -5.50$, followed by the Scene condition, $\hat{\beta} = -0.68°$, $t = -2.81$. See Table 1 for details.

## Saccadic dead time

SDT may be defined as the last point in time at which novel stimulus information may be incorporated by the system responsible for preparing a saccadic response. Thus, SDT may also be described as the point of no return in the preparation of a saccade. Once the point of no return in saccade programming has been

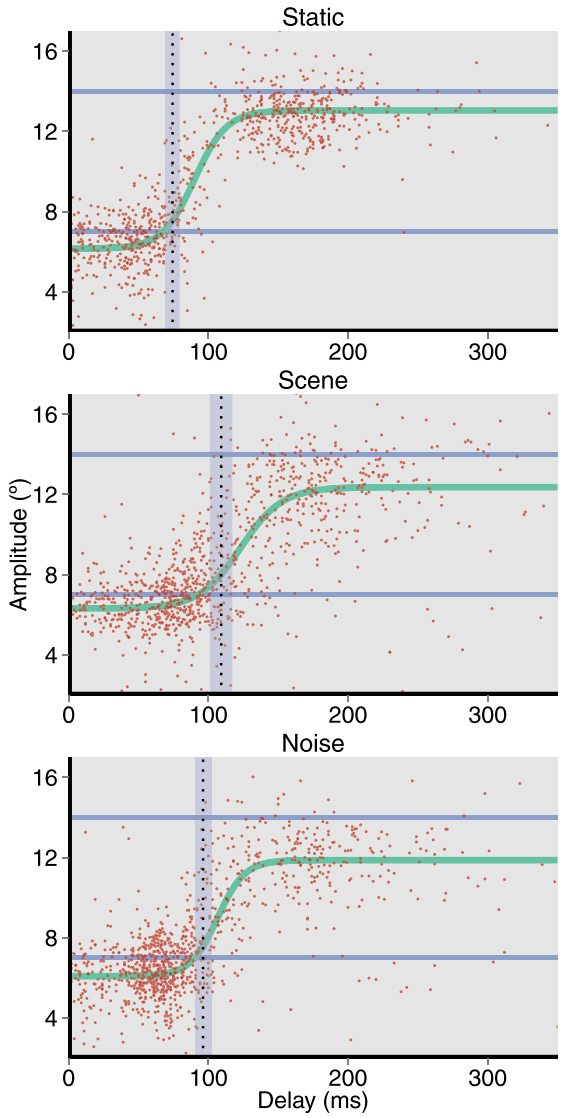


Figure 3. Amplitude Transition Functions for the three experimental conditions in Experiment 1. Points represent the amplitudes of saccades initiated in response to the presentation of double-step targets. On the *x*-axis, the delay (D) represents the amount of time elapsed between the presentation of the second target and the onset of the saccade. The *y*-axis represents the amplitude of the resulting saccade. The green lines represent the best-fitting fixed-effects curves estimated with a nonlinear mixed-effects regression. The horizontal solid blue lines represent the physical distance of the targets in relation to fixation location (7° and 14°). The vertical dotted black lines represent the mean SDT in each condition derived from the by-participant SDT estimates. The light blue bands surrounding the means are bootstrapped confidence intervals.

reached, that saccade may no longer be modified or canceled.

To estimate the SDT from the fitted ATFs, we determined the largest value of D for which the amplitude of the response saccade showed no evidence of incorporating the second target position into the response. To define saccade amplitudes in the double-step condition that are inconsistent with responses programmed to the first target location, we used the distribution of response amplitudes observed in the single-step condition. In the single-step condition, saccades target the first location and are characterized by a distribution of amplitudes that cluster near the first target location. For each participant and each condition, we measured the amplitude that corresponds to the 95th percentile of responses. Such a value provides a cutoff for amplitudes that are rarely associated with responses to the initial target location. This cutoff was used to define an amplitude threshold in the double-step condition such that responses with an amplitude beyond this point were considered to be influenced by the second target step.[1]

| Participant | Static | Scene | Noise |
|---|---|---|---|
| 1 | 79 | 115 | 94 |
| 2 | 57 | 127 | 95 |
| 3 | 67 | 103 | 83 |
| 4 | 75 | 89 | 89 |
| 5 | 73 | 105 | 96 |
| 6 | 83 | 111 | 100 |
| 7 | 85 | 115 | 106 |
| 8 | 79 | 112 | 100 |
| 9 | 69 | 115 | 92 |
| 10 | 83 | 118 | 114 |
| 11 | 86 | 134 | 115 |
| 12 | 70 | 90 | 85 |
| 13 | 65 | 90 | 85 |
| 14 | 84 | 122 | 104 |
| 15 | 65 | 95 | 85 |
| Mean | 74 | 109 | 96 |

Table 2. Estimated Saccadic Dead Time (ms) in Experiment 1.

As the ATF was estimated with a nonlinear mixed-effects regression with by-participant random intercepts and slopes of experimental condition, it was possible to provide an independent estimate of SDT for each participant in each condition. Specifically, estimates of SDT were computed for individual participants by inspecting the responses predicted by the individual-level data from the mixed-effects regression. Individual-level SDT estimates, along with mean SDT in the three conditions, are provided in Table 2. To statistically validate the differences in observed means, we used a bootstrap procedure to estimate the distribution of mean SDT in each of the three conditions (Efron & Tibshirani, 1993). In summary, the SDT observations recorded in each condition were resampled with replacement 10,000 times, and for each iteration, a sample mean was recorded. Through this method, we constructed a distribution of sample means based on bootstrapped data. These distributions were then used to construct 95% confidence intervals (CIs)

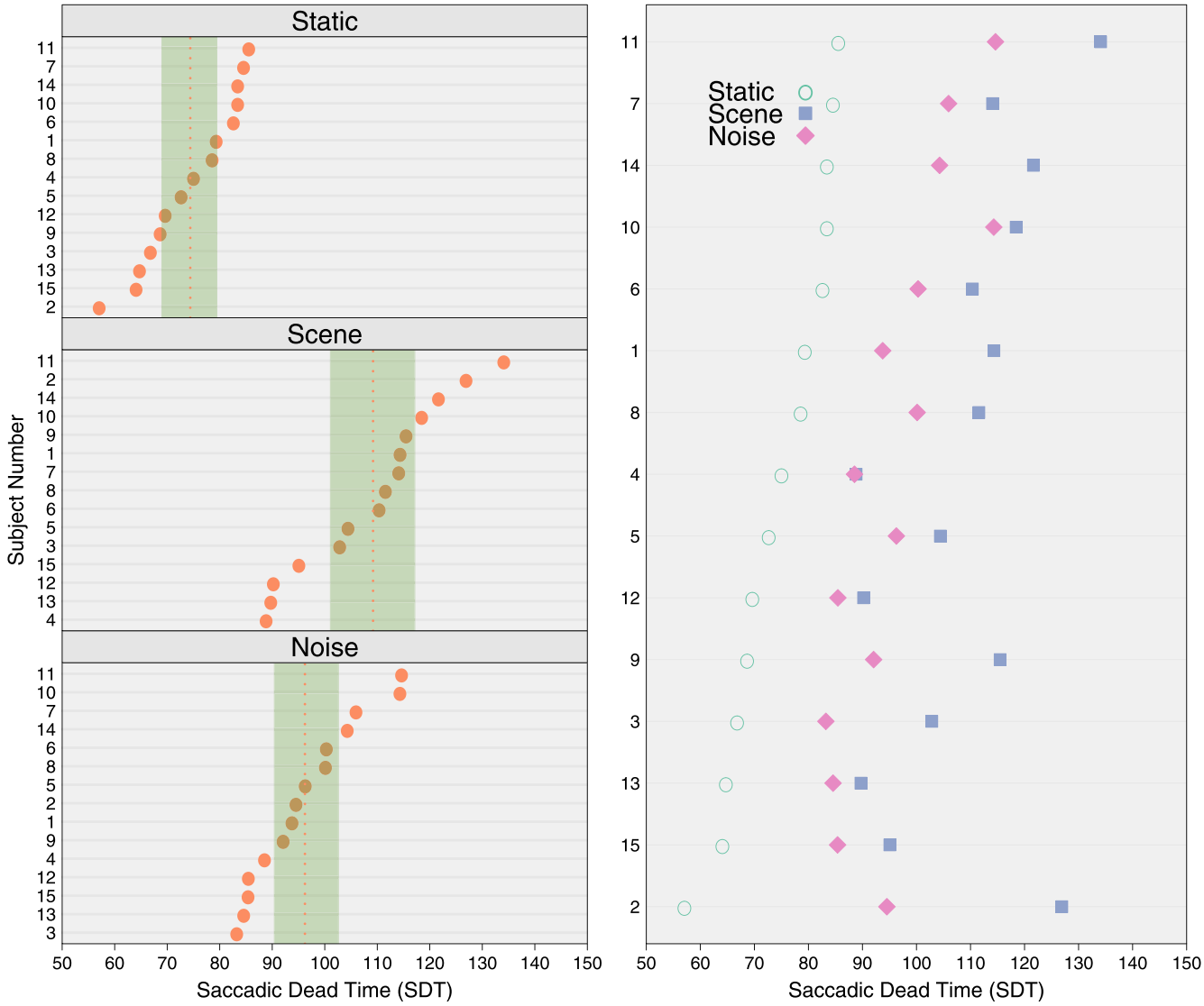


Figure 4. Saccadic Dead Time (SDT) estimates (ms) for the three experimental conditions in Experiment 1. The left plot shows by-participant estimates of SDT. For a given condition, from bottom to top, values are ordered from lowest to highest. The vertical dotted lines represent the mean value of SDT in each condition. The green band surrounding the mean is the bootstrapped confidence intervals of the mean estimate. The right plot shows participants ordered by the magnitude of their estimated SDT value in the Static condition.

around the observed means. We defined a significant difference between means in any two conditions as occurring when the bootstrapped CIs did not overlap. To control for familywise error rates, Bonferroni correction was applied to the CIs. The mean SDTs and CIs were derived from the fitted model. The mean SDT was shortest in the Static condition ($M = 74.35$ ms, $CI_{low} = 69.03$, $CI_{high} = 79.65$), followed by the Noise condition ($M = 96.24$ ms, $CI_{low} = 90.49$, $CI_{high} = 102.69$) and the Scene condition ($M = 109.18$ ms, $CI_{low} = 100.94$, $CI_{high} = 117.27$). Therefore, according to the bootstrapped hypothesis-testing procedure, statistically reliable differences were observed between SDT in the Static versus Scene and Static versus Noise conditions. Furthermore, a trend was observed in the Scene versus Noise conditions such that SDT tended to be longer in the Scene than in the Noise condition. Inspection of the individual measures of SDT (see Figure 4) shows that the SDT for each participant was numerically larger in the Scene than in the Noise condition.

## Additional analyses

Additional analyses explored whether the structural differences in the stimulus content in the two dynamic tasks were associated with differences in global eye-movement parameters. To compare mean saccade amplitudes and fixation durations in the Noise and Scene conditions, only those saccades that were not

generated in response to a target step were considered. For saccade amplitudes, no statistically significant difference was found ($p > 0.05$). Saccade amplitudes to single-step targets were also analyzed. They were shortest in the Noise condition ($M = 6.20°$), followed by the Static condition ($M = 6.44°$) and the Scene conditions ($M = 6.51°$). Pairwise $t$ tests with a family-wise error rate of 0.01 were used to statistically validate the mean differences. The difference between Scene and Noise was statistically significant, $t(14) = 3.49$, $p = 0.004$, as was the difference between Static and Noise, $t(14) = 2.80$, $p = 0.01$. The difference between Static and Scene did not reach statistical significance.

Mean fixation durations were longer in the Noise condition ($M = 324$ ms) compared with the Scene condition ($M = 276$ ms), and this difference was significant, $t(14) = 4.28$, $p = 0.0008$. This pattern of results replicates a previous study that compared $1/f$ filtered scenes with naturalistic scenes (Kaspar & König, 2011).

In addition, we analyzed saccadic reaction times to single-step stimuli. Saccadic responses were fastest in the Noise condition ($M = 162$ ms), followed by the Scene ($M = 177$ ms) and Static ($M = 195$ ms) conditions. Saccadic response times were significantly longer in the Scene than in the Noise condition, $t(14) = 2.51$, $p = 0.03$. The difference between the Static and the Noise condition was also significant, $t(14) = 4.19$, $p = 0.001$. The difference between the Static and the Scene and fitted ATFs are shown in Figure 5.

## Experiment 2

A second experiment was conducted to compare SDT under conditions in which the structure of the background is varied but in which participants do not explore the scene prior to target onset. In this experiment, participants conducted three versions of the Static task from Experiment 1. In the first condition, the task is conducted on a uniformly black background. We call this condition the Uniform condition and note that it is identical to the Static condition in Experiment 1. In the second condition (Scene condition), the same task is conducted, but the background is replaced by one of the naturalistic scenes used in Experiment 1. In a final condition, the task is conducted with $1/f$ stimuli presented as the background.

As movement is controlled for in this study, any observed difference in SDT can be more directly interpreted to result from differences in the structure of the background. Specifically, we predict that if increases in SDT observed in Experiment 1 are due to the additional structure of the backgrounds in the Scene and Noise conditions, then the structured background conditions in Experiment 2 should reveal an elevated SDT relative to the uniform background.

The experiment was conducted on an additional seven male (including one author, R.C.W.) and five female participants with an average age of 24 years who did not participate in Experiment 1. The three conditions in Experiment 2 (Uniform, Scene, Noise) paralleled in nearly all details the Static condition from Experiment 1. The primary difference was that in two of the conditions, the backgrounds were replaced by scenes (Scene conditions) or phase noise images (Noise condition) that were used in Experiment 1. On each trial, the specific scene or noise image was randomly selected from the set of stimuli used in Experiment 1. Participants were explicitly told that the scene or noise background content was not relevant to the completion of their task.

### Results

The approach to the analysis of the results in Experiment 2 was conducted in an analogous way to the Static condition in Experiment 1. The nonlinear mixed-effects model in Experiment 2 included the same random effects structure as the model in Experiment 1. That is, by-participant random intercepts and slopes were included for $\alpha$, $\beta$, $\gamma$, and $\delta$. The model estimates are summarized in Table 3.

The estimated SDT in the uniform condition was $M_{Uniform} = 71.39$ ms, $CI_{low} = 65.00$, $CI_{high} = 78.26$. In the Scene condition, SDT was estimated at $M_{Scene} = 80.70$ ms, $CI_{low} = 75.50$, $CI_{high} = 86.11$ and was estimated at $M_{Noise} = 70.83$ ms, $CI_{low} = 65.93$, $CI_{high} = 76.55$ in the Noise condition. As was introduced in the analysis of Experiment 1, lack of overlap in Bonferroni-corrected CIs was used as a criterion for rejecting the null hypothesis. Accordingly, there was no difference in SDT between the Uniform and Noise conditions, as is evidenced by the strongly overlapping CIs (Figure 6). The CI for the Scene condition slightly overlapped with the other two CIs. However, inspection of the pattern of results at the level of individual SDT responses revealed a strong trend for longer SDT in the Scene condition when compared with the uniform and phase conditions. For 10 of the 12 participants, the longest SDT response was observed in the Scene condition (Figure 6; Table 4).

We also analyzed saccade latencies on trials in which only the initial target was presented. Latency was lowest in the uniform condition ($M = 203$ ms), followed by the Noise condition ($M = 208$ ms) and Scene condition ($M = 214$ ms). The differences between Scene versus Noise and Noise versus Uniform did not reach statistical significance, but the difference between Scene

| | Fixed effects | | | Random effects | |
|---|---|---|---|---|---|
| Parameter | Estimate | SE | *t* value | Parameter | $\sigma$ |
| Lower bound ($\alpha$) | | | | | |
| Uniform (Intercept) | 6.55 | 0.15 | 44.15 | Uniform | 0.42 |
| Noise | 0.37 | 0.13 | 2.96 | Noise | 0.10 |
| Scene | 0.35 | 0.12 | 2.92 | Scene | 0.02 |
| Upper bound ($\beta$) | | | | | |
| Uniform (Intercept) | 13.06 | 0.16 | 79.48 | Uniform | 0.52 |
| Noise | −0.08 | 0.14 | −0.57 | Noise | 0.33 |
| Scene | −0.10 | 0.13 | −0.74 | Scene | 0.34 |
| Inflection ($\delta$) | | | | | |
| Uniform (Intercept) | 86.62 | 3.79 | 22.88 | Uniform | 12.63 |
| Noise | 2.71 | 2.90 | 0.94 | Noise | 8.63 |
| Scene | 9.96 | 2.31 | 4.32 | Scene | 6.34 |
| Scale ($\gamma$) | | | | | |
| Uniform (Intercept) | 0.10 | 0.01 | 7.83 | Uniform | 0.04 |
| Noise | −0.01 | 0.01 | −1.04 | Noise | 0.02 |
| Scene | −0.004 | 0.02 | −0.22 | Scene | 0.04 |
| Error term | | | | | |
| ε | | | | | 1.11 |

Table 3. ATF regression model (Experiment 2). *Notes:* Estimated parameters for the four-parameter logistic regression model in Experiment 2. Means, standard errors, and *t* values of fixed effects; standard deviations of the random effects.

and Uniform was statistically significant, $t(11) = 0.04$, $p < 0.04$.

Saccade amplitudes to single-step targets were also analyzed. They were shortest in the uniform condition ($M = 6.83°$), followed by the Scene condition ($M = 6.99°$) and the Noise condition ($M = 7.01°$). The difference between the Uniform and Scene conditions was statistically significant, $t(11) = -2.57$, $p = 0.03$, as was the difference between Uniform and Noise, $t(11) = -3.29$, $p = 0.007$. The difference between Scene and Noise was not statistically significant.

## Discussion

The goal of this study was to extend results regarding the programming of saccadic responses to briefly presented targets from a static context to a dynamic context that more closely resembles the environment that saccades are programmed in when viewing naturalistic scenes. In the three conditions of the main experiment (Static, Scene, and Noise), single- or double-step targets were presented while participants were engaged in a stable fixation. In the Static task, participants responded to targets following an extended period of fixation on a central cross. During both the

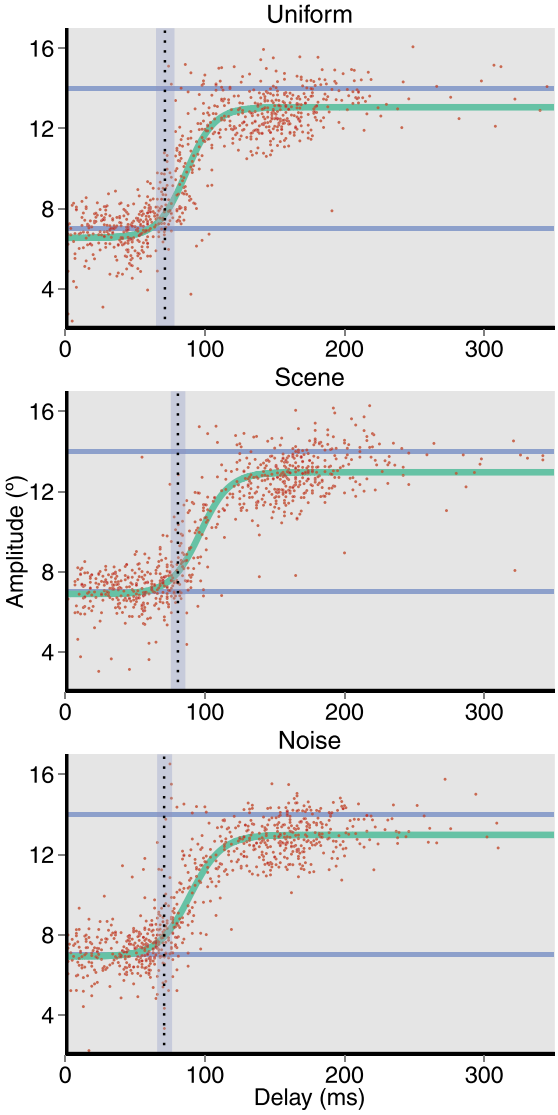


Figure 5. Amplitude transition functions for the three experimental conditions in Experiment 2. Points represent the amplitude of saccades initiated in response to the presentation of double-step targets. On the *x*-axis, the delay (D) represents the amount of time elapsed between the presentation of the second target and the onset of the saccade. The *y*-axis represents the amplitude of the resulting saccade. The green lines represent the best-fitting fixed-effects curves estimated with a nonlinear mixed-effects regression. The horizontal solid blue lines represent the physical distance of the targets in relation to fixation location (7° and 14°). The vertical dotted black lines represent the mean SDT in each condition derived from the by-participant SDT estimates. The light blue bands surrounding the means are bootstrapped confidence intervals.

Scene and Noise conditions, participants responded to targets that were presented at the onset of a fixation made while exploring the image. The double-step logic used by Becker and Jürgens (1979) was ported to the present study, which allowed comparative inferences to

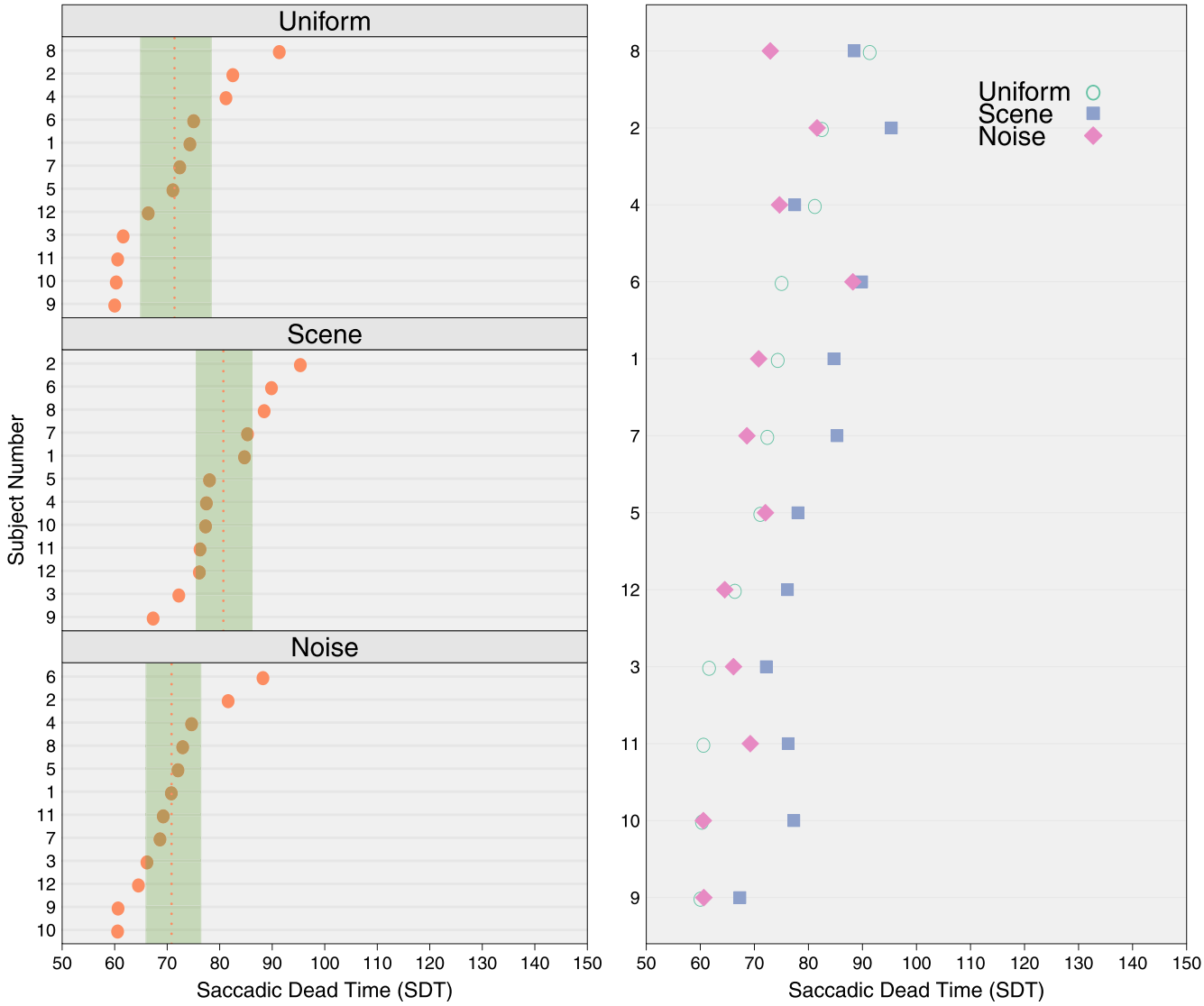


Figure 6. Saccadic dead time (SDT) estimates (ms) for the three experimental conditions in Experiment 2. The left plot shows by-participant estimates of SDT. For a given condition, from bottom to top, values are ordered from lowest to highest. The vertical dotted lines represent the mean value of SDT in each condition. The green band surrounding the mean is the bootstrapped confidence intervals of the mean estimate. The right plot shows participants ordered by the magnitude of their estimated SDT value in the uniform condition.

be made regarding the time course of saccade programming in the three conditions.

Modification of saccade programming timelines has been an influential assumption used by models that attempt to describe eye-movement control in both scene viewing and in reading (Reichle et al., 1998; Engbert et al., 2005; Nuthmann et al., 2010; Trukenbrod & Engbert, 2014). Previously, these assumptions have primarily been warranted by double-step experiments using tasks similar to the Static condition implemented in our Experiment 1. In the present study, we observed the characteristic logistic shape of the ATF in both static and dynamic viewing conditions, which provides an important confirmation that reprogramming mechanisms that operate in static contexts operate in an analogous manner within a dynamic scene-viewing context.

ATFs permit inferences to be made regarding the time course of saccade preparation. SDT represents the point in time prior to the onset of a saccade at which that saccade may no longer be modified by updated visual information. In Experiment 1, we found that when double-step targets are presented in a static movement context in which the targets are presented on a black background, SDT is lower than it is when targets are presented in a dynamic double-step context overlaid on structured backgrounds. This difference in SDT is also complemented by a strong trend toward observing longer SDT in the Scene as compared with the Noise condition, suggesting a possible influence of

| Participant | Uniform | Scene | Noise |
|---|---|---|---|
| 1 | 74 | 85 | 71 |
| 2 | 82 | 95 | 82 |
| 3 | 62 | 72 | 66 |
| 4 | 81 | 77 | 75 |
| 5 | 71 | 78 | 72 |
| 6 | 75 | 90 | 88 |
| 7 | 72 | 85 | 69 |
| 8 | 91 | 88 | 73 |
| 9 | 60 | 67 | 61 |
| 10 | 60 | 77 | 61 |
| 11 | 60 | 76 | 69 |
| 12 | 66 | 76 | 65 |
| Mean | 71 | 81 | 71 |

Table 4. Estimated saccadic dead time (ms) in Experiment 2.

scene content per se. In Experiment 2, we isolated the influence of background on SDT by comparing three static movement conditions, which differed in the structure of the background stimulus. We found that SDT was no different in the Uniform and Noise conditions, but there was a tendency to observe a larger SDT in the Scene condition. Furthermore, by removing movement from the task, SDT estimates in the Scene and Noise conditions were greatly reduced in comparison to Experiment 1.

In Experiment 1, a notable difference between the Static and Scene task is that within the Scene task, but not the Static task, the targets are presented during a dynamic movement context. A further difference between these tasks is the high-level cognitive processes that are assumed to be activated during scene exploration. In the Scene task, participants are under instructions to view the stimulus to prepare for a later recall phase of the experiment. As a result, in the Scene condition, participants are engaged in a more complex task that is more engaging to higher-level cognition. Therefore, it might be predicted that higher-order operations specific to the processing of scene elements are implicated in the elevated SDT that was observed in the Scene condition relative to the Static condition. However, a direct comparison between these two conditions does not distinguish between (a) influences that are due to the dynamic movement and (b) differences due to additional high-level scene structure. The Noise condition in Experiment 1 was explicitly included to discriminate between these two alternatives. If scene structure does play a role in determining SDT, then it would be expected that SDT in the Scene condition is also elevated relative to the Noise condition. The pattern of results observed in Experiment 1 suggests that there may be a unique influence of scene content, as a trend toward longer SDT was observed for scene backgrounds. SDT was estimated to be 13 ms longer in the Scene compared with the Noise condition. Although this difference did not reach statistical significance, the pattern of individual results showed a clear trend toward increased SDT in the Scene as compared with the Noise condition. Results from Experiment 2 also support the view that scene content may play a role in determining SDT. As in Experiment 1, a trend for longer SDT was observed in the Scene condition relative to the non-Scene conditions. This influence on SDT occurred despite participants receiving instructions that the scene content was not relevant in any way to the successful completion of their task.

The static versus dynamic movement context also appears to play a relatively strong role in determining SDT. In Experiment 1, we observed a strong effect of movement on SDT. SDT in the two movement conditions (Scene and Noise) was elevated from the Static condition by 22 and 35 ms, respectively. As was previously argued, the SDT increase in the Noise condition relative to the Static condition should not be considered to arise from the additional structure present in the $1/f$ amplitude spectra. Experiment 2 provides support for this difference being one that is linked to movement. When movement was controlled for, no difference between the Static (Uniform) and Noise conditions was observed, and the SDT magnitudes in the Scene and Noise conditions were reduced. This was in contrast to Experiment 1 in which these backgrounds led to observed differences in SDT. This suggests that differences in SDT can arise when the eyes are in motion and actively engaged in a task, as compared to when they are at rest, waiting for stimulus presentation.

We suggest that the difference in SDT between the static and dynamic tasks arises partially due to a form of saccadic prepreparation that occurs within the Static task but is less likely to be active within the dynamic tasks. We speculate that by presenting the first target, initial stages of saccade preparation may be initiated both to the location of the visible target and to the location at which the second target is predicted to appear. In the Static task, the location of the second target is highly predictable as participants are always fixating the same location (central cross) when it appears. In contrast, estimates of the location of the target position in the dynamic task are likely to be far more variable in that there is limited evidence on which to base such predictions. As a result of such predictability, we suggest that the saccade motor system may have primed the saccade to the second target prior to the target onset. An alternative possibility is that SDT is elevated in the dynamic tasks due to the occurrence of multiple saccade plan modifications. In this account, when the first target is presented, a modification to the current saccade plan is initiated. Then, upon presentation of the second target, this updated saccade plan

must receive additional modification to reach the final target position. It may be the case that such cascaded modifications require the presence of increased SDT. As current models of saccade programming do not address such a scenario, it is difficult to make specific predictions. However, such a simulation would be an interesting and worthwhile exercise.

Stochastic process models of double-step performance in simple tasks provide an elegant framework with which to interpret how these hypothesized differences between tasks could result in SDT variability. Camalier et al. (2007) used a race model architecture to model double-step performance. In this model, saccades are generated as a consequence of a race between a saccade program targeting the first location and an additional stochastic cancellation process that is initiated upon presentation of the second target. The quantity known as the target step reaction time (TSRT) expresses the amount of time that is required by the cancellation process to successfully inhibit saccades to the first target. As lower values of TSRT correspond to lower SDT, the race model predicts that a process that results in a speed up of the cancellation process will also result in a reduction in SDT. The authors applied the race model logic to a double-step task in which the second target step competed with distractor stimuli for saccade target selection. They demonstrated that TSRT was longer when the target and distractors were more similar in color, illustrating that SDT is likely to be longer in conditions where acquisition of the second target is made more difficult. Therefore, it may be the case that lower SDT in the Static task may result from a decrease in the amount of time required for the cancellation process to complete because of the predictable characteristics of the second target location.

Ludwig et al. (2007) measured SDT in two double-step tasks. In the first task, a gap paradigm (Saslow, 1967) was used to observe SDT under conditions in which the latency of responses to targets is systematically varied. Saccade latency was manipulated by completely removing the central fixation cross prior to the onset of the saccadic response. Despite the fact that, on gap trials, saccade latency was considerably reduced (by 28%), no difference in SDT was observed. In a second static double-step experiment, the authors manipulated the angle of separation between two double-step targets such that the angle between the two targets on a given trial varied between 30° and 90°. In this task, they found a positive relationship between the degree of separation and the estimated value of SDT. They reported a mean SDT of 65 ms at the smallest separation of 30° and an SDT of 113 ms at the largest separation of 120°. The authors introduced a population coding account to explain the pattern of results. In this account, when the initial and final target are presented in close spatial proximity, there is shared activation in direction coded movement neurons for the initial and final targets. Therefore, presentation of the initial target provides a form of prepreparation for the final target movement. Because of such prepreparation, neurons coding for movement toward the second target have been partially activated and therefore require less time to reach the threshold to initiate a saccadic response. This population coding argument has been shown to adequately account for the effect of target angle separation on SDT (Ludwig et al., 2007). However, the population coding account as formulated in Ludwig et al. (2007) cannot directly account for the results reported here. In all double-step tasks, the angle between the target stimuli was maintained to be a constant 0°. Therefore, according to this framework, activation of neurons coding movement to the second target location by presentation of the first target should be equivalent in both static and dynamic tasks.

The estimated values for SDT observed in our study (Static: 74 ms, Scene: 109 ms, Noise: 96 ms) also provide a measure of empirical grounding to the values for the duration of the nonlabile stage of saccade programming that have previously been suggested (see Table 5). It is important to note that the concept of SDT does not map directly onto that of a nonlabile stage. First, the magnitude of the SDT estimates may vary depending on the method chosen to estimate the point of no return. Although a consistent method may be used to compare SDT within a single study, the magnitude of the SDT estimates may vary depending on the amplitude cutoff chosen to calculate SDT. Furthermore, any SDT estimate that is derived from double-step response data includes delays in transmission of the visual information to regions of the brain responsible for saccadic decisions. Because the nonlabile stage of saccade programming is conceptualized to operate independently of such transmission delays, deriving a nonlabile estimate from SDT requires a subtraction of this afferent delay. Neurophysiological evidence suggests that afferent delays are on the order of 50 ms (for review, see Reichle & Reingold, 2013).

Given the current method of calculating SDT, taking such delays into account would suggest that in scene viewing the duration of the nonlabile stage is approximately 60 ms. Because of the task-dependent nature of SDT, we believe that it is difficult to generalize this result to other domains. We suggest that an intriguing possibility for extension of this work would be to compare directly SDT estimates across both reading and scene-viewing tasks, thereby permitting more concrete inferences regarding the task-dependent nature of SDT. Further comment is warranted regarding the efferent delays that are known to exist in the transmission of oculomotor decisions from the brain to the eye. Such delays are thought to be on the order of

| Model | | Primary task | Duration (ms) | Other task | Duration (ms) |
|---|---|---|---|---|---|
| E-Z Reader | | | | | |
| 1-5 | Reichle et al. (1998) | Reading | 50 | — | — |
| 9 | Pollatsek et al. (2006) | Reading | 25 | — | — |
| 10 | Reichle et al. (2012) | Reading | 25 | Search | 25 |
| 10 | Reichle et al. (2012) | Reading | 25 | Z-String | 25 |
| SWIFT | | | | | |
| I | Engbert, Longtin, and Kliegl (2002) | Reading | 41.6 | — | — |
| II | Engbert et al. (2005) | Reading | [6.1, 58.7] | — | — |
| II | Nuthmann and Engbert (2009) | Reading | [6.1, 58.7] | Z-String | [20.5, 59.6] |
| III | Schad and Engbert (2012) | Reading | [50, 51.3] | Shuffled | [50, 51.3] |
| CRISP | | | | | |
| | Nuthmann et al. (2010) | Scene viewing | 40 | — | — |
| | Nuthmann and Henderson (2012) | Scene viewing | 40 | Reading | 14 |
| ICAT | | | | | |
| | Trukenbrod and Engbert (2014) | Search | 40 | Reading | 40 |

Table 5. Comparison of model nonlabile estimates. *Notes:* As of SWIFT-II, the model predicts a continuum of nonlabile durations. A range of values are provided that reflect short (1°) and long (10°) reading saccades.

20 ms (Becker, 1991). In the saccade programming architecture of the models that we have discussed, such delays must logically occur toward the latter stages of the nonlabile stage. The duration of the efferent delay places a lower bound on the duration of the nonlabile stage as no stimulus information may contribute to the programming of a saccadic response after the movement signal has been sent to the motor effectors.

The present results are particularly informative for models of scene perception (Nuthmann et al., 2010), reading (Engbert et al., 2005; Reichle et al., 1998), and visual search (Trukenbrod & Engbert, 2014) that use multistage saccade programming architectures. Table 5 shows the parameters used to specify the mean duration of the nonlabile stage of saccade programming in these models. Surprisingly, there is little consistency with regard to the duration of the nonlabile stage within a given task. In some cases, the models predict durations that are less than the minimum interval required by the efferent delay. This represents a potential difficulty in that such models aim to synthesize empirical knowledge regarding the timeline of oculomotor control into predictions about behaviors in high-level tasks. One of the contributions of the present study is to provide a principled experimental grounding for assumptions that are critical to such behavioral predictions. However, it is important to note that differences between the double-step tasks under investigation here and other typical scene-viewing contexts do exist. For example, in the dynamic viewing tasks, participants are aware that on each trial, a sudden target onset will occur after the scene has been explored for some number of fixations. Consequently, it is possible that toward the later stages of a trial, participants may begin to modify their viewing strategies. It is therefore worthwhile considering that these tasks are designed to study saccade programming during scene perception tasks that include a target acquisition component.

There is also some discord between models on the question of task-dependent differences in the duration of the nonlabile stage. In both the E-Z Reader model (Reichle et al., 2012) and the CRISP model (Nuthmann et al., 2010), there is no mechanism provided by which the mean nonlabile duration may vary within a specific task.[2] In Reichle et al. (2012), an estimated mean duration of 25 ms remained constant both within and across reading, z-string reading, and search tasks. In simulations with the CRISP model (Nuthmann & Henderson, 2012), a comparison was made between model predictions of fixation durations in reading and in scene viewing. In these simulations, the duration of the nonlabile stage was permitted to vary across tasks (reading: 14 ms; scene viewing: 40 ms). Similarly to the CRISP model, the SWIFT model also predicts task-dependent differences in mean nonlabile durations (Nuthmann & Engbert, 2009). However, as of SWIFT-II, the model additionally assumes that nonlabile durations vary systematically within a task by scaling the nonlabile duration to the length of the resulting saccade (Engbert et al., 2005). SWIFT-II was evaluated with data from the Potsdam Sentence Corpus, in which one letter subtends 0.38° and/or 0.45° of visual angle (Nuthmann & Kliegl, 2009). For short saccades (≈ 1°) SWIFT-II predicts average nonlabile durations as long as 58.7 ms. At the other end of the continuum, long saccades (≈ 10°) are predicted to have an average nonlabile duration of 6.1 ms. Therefore, between-task differences in average saccade amplitude may translate into differences in nonlabile durations. However, it should also be noted that estimated parameters in a later version of the SWIFT model result in a model with no between- or within-task variability in nonlabile dura-

tions (Schad & Engbert, 2012). A recent model, ICAT (Trukenbrod & Engbert, 2014), has modeled fixation durations in both visual search and reading tasks. In ICAT, a fixed nonlabile duration of 40 ms was assumed for both tasks. Future modeling efforts would greatly benefit from addressing the question of what assumptions are being made with regard to the nonlabile values that are selected or estimated from the data.

## Conclusions

Computational models have played an important role in our understanding of oculomotor control behavior in high-level tasks such as scene viewing and reading. However, these models have often relied on simple stimuli used in basic psychophysical paradigms to inform the development of their architectures. Here, using a gaze-contingent double-step paradigm, we have generalized these findings to a more ecologically valid context and revealed important task differences in saccade programming. The present results are particularly informative for the understanding of saccade programming during scene viewing. However, we suggest that future work should directly investigate the processes of saccade cancellation in reading and other task contexts. These studies would provide further generalization for a role of saccade cancellation in eye-movement control, as well as provide empirical validation for task-specific modeling efforts.

*Keywords: saccade programming, scene viewing, double-step, eye-movement control, reading*

## Acknowledgments

Commercial relationships: none.
Corresponding author: R. Calen Walshe.
Email: calen.walshe@gmail.com.
Address: Psychology Department, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK.

## Footnotes

[1] We thank Casimir Ludwig for suggesting this method. We also note that calculating SDT in such a manner resulted in estimates that were increased in comparison to an alternative method (see Ludwig et al., 2007).

[2] Although there is no systematic variability built into the nonlabile duration, stochastic variability does enter these models via the duration being sampled from a gamma distribution.

## References

Aslin, R. N., & Shea, S. L. (1987). The amplitude and angle of saccades to double-step target displacements. *Vision Research, 27*, 1925–1942.

Becker, W. (1991). Saccades. In R. H. S. Carpenter (Ed.), *Eye-movements: Vol. 8 in vision and visual dysfunction* (pp. 95–137). Boston, MA: CRC Press.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research, 19*, 967–983.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433–436.

Camalier, C. R., Gotler, A., Murthy, A., Thompson, K. G., Logan, G. D., Palmeri, T. J., & Schall, J. D. (2007). Dynamics of saccade target selection: Race model analysis of double step and search step saccade production in human and macaque. *Vision Research, 47*, 2187–2211.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers, 34*, 613–617.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman and Hall.

Einhäuser, W., Rutishauser, U., Frady, E. P., Nadler, S., König, P., & Koch, C. (2006). The relation of phase noise and luminance contrast to overt attention in complex visual stimuli. *Journal of Vision, 6*(11):1, 1148–1158, http://www.journalofvision.org/content/6/11/1, doi:10.1167/6.11.1. [PubMed] [Article]

Einhäuser, W., Rutishauser, U., & Koch, C. (2008). Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *Journal of Vision, 8*(2):2, 1–19, http://www.journalofvision.org/content/8/2/2, doi:10.1167/8.2.2. [PubMed] [Article]

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42*, 621–636.

Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*, 777–813.

Findlay, J. M., & Harris, L. R. (1984). Small saccades to double-stepped targets moving in two dimensions. In A. G. Gale & F. Johnson (Eds.), *Theoretical and applied aspects of eye movement research* (pp. 71–78). Amsterdam: Elsevier.

Foulsham, T., Kingstone, A., & Underwood, G. (2008). Turning the world around: Patterns in saccade direction vary with picture orientation. *Vision Research, 48*, 1777–1790.

Henderson, J. M., Nuthmann, A., & Luke, S. G. (2013). Eye movement control during scene viewing: Immediate effects of scene luminance on fixation durations. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 318–322.

Henderson, J. M., & Pierce, G. L. (2008). Eye movements during scene viewing: Evidence for mixed control of fixation durations. *Psychonomic Bulletin & Review, 15*, 566–573.

Ho-Phuoc, T., Guyader, N., Landragin, F., & Guérin-Dugué, A. (2012). When viewing natural scenes, do abnormal colors impact on spatial or temporal parameters of eye movements? *Journal of Vision, 12*(2):4, 1–13, http://www.journalofvision.org/content/12/2/4, doi:10.1167/12.2.4. [PubMed] [Article]

Kaspar, K., & König, P. (2011). Overt attention and context factors: the impact of repeated presentations, image type, and individual motivation. *PLoS One, 6*, e21719.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception, 36*, 14.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research, 41*, 3559–3565.

Ludwig, C. J. H., Mildinhall, J. W., & Gilchrist, I. D. (2007). A population coding account for systematic variation in saccadic dead time. *Journal of Neurophysiology, 97*, 795–805.

Mills, M., Hollingworth, A., Van der Stigchel, S., Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. *Journal of Vision, 11*(8):17, 1–15, http://www.journalofvision.org/content/11/8/17, doi:10.1167/11.8.17. [PubMed] [Article]

Nuthmann, A., & Einhäuser, W. (2015). A new approach to modeling the influence of image features on fixation selection in scenes. *Annals of the New York Academy of Sciences, 1339*, 82–96.

Nuthmann, A., & Engbert, R. (2009). Mindless reading revisited: An analysis based on the SWIFT model of eye-movement control. *Vision Research, 49*, 322–336.

Nuthmann, A., & Henderson, J. M. (2010). Object-based attentional selection in scene viewing. *Journal of Vision, 10*(8):20, 1–19, http://www.journalofvision.org/content/10/8/20, doi:10.1167/10.8.20. [PubMed] [Article]

Nuthmann, A., & Henderson, J. (2012). Using CRISP to model global characteristics of fixation durations in scene viewing and reading with a common mechanism. *Visual Cognition, 20*, 457–494.

Nuthmann, A., & Kliegl, R. (2009). An examination of binocular reading fixations based on sentence corpus data. *Journal of Vision, 9*(5):31, 1–28, http://www.journalofvision.org/content/9/5/31, doi:10.1167/9.5.31. [PubMed] [Article]

Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational model of fixation durations in scene viewing. *Psychological Review, 117*, 382–405.

Pinheiro, J., & Bates, D. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2014). *nlme: Linear and nonlinear mixed effects models*. R package version 3.1–117. Vienna, Austria: R Foundation for Statistical Computing.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52*, 1–56.

R Development Core Team. (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ray, S., Schall, J. D., & Murthy, A. (2004). Programming of double-step saccade sequences: Modulation by cognitive control. *Vision Research, 44*, 2707–2718.

Reichle, E., Pollatsek, A., Fisher, D., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*, 125–157.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2012). Using E-Z Reader to simulate eye movements in nonreading tasks: A unified framework for understanding the eye-mind link. *Psychological Review, 119*, 155–185.

Reichle, E. D., & Reingold, E. M. (2013). Neurophysiological constraints on the eye-mind link. *Frontiers in Human Neuroscience, 7*, 1–6.

Reingold, E. M., Reichle, E. D., Glaholt, M. G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from a survival analysis of fixation durations. *Cognitive Psychology, 65*, 177–206.

Saslow, M. (1967). Effects of components of displacement-step stimuli upon latency for saccadic eye movement. *Journal of the Optical Society of America, 57*, 1024–1029.

Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model. *Visual Cognition, 20*, 391–421.

Shioiri, S. (1993). Postsaccadic processing of the retinal image during picture scanning. *Perception & Psychophysics, 53*, 305–314.

Smith, T. J., & Mital, P. K. (2013). Attentional synchrony and the influence of viewing task on gaze behavior in static and dynamic scenes. *Journal of Vision, 13*(8):16, 1–24, http://www.journalofvision.org/content/13/8/16, doi:10.1167/13.8.16. [PubMed] [Article]

SR Research Ltd. (2006). *EyeLink programmer's guide*. (3.0 ed.). Retrieved from download.sr-support.com/dispdoc.

Tatler, B. W., & Vincent, B. T. (2008). Systematic tendencies in scene viewing. *Journal of Eye Movement Research, 2*(2):5, 1–18.

Trukenbrod, H. A., & Engbert, R. (2014). ICAT: A computational model for the adaptive control of fixation durations. *Psychonomic Bulletin & Review, 21*, 904–934.

Walshe, R. C., & Nuthmann, A. (2013). Programming of saccades to double-step targets in scene viewing: A test of assumptions present in the CRISP model. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 1569–1574). Berlin. Cognitive Science Society.

Walshe, R. C., & Nuthmann, A. (2014). Asymmetrical control of fixation durations in scene viewing. *Vision Research, 100*, 38–46.

Westheimer, G. (1954). Eye movement responses to a horizontally moving visual stimulus. *Archives of Ophthalmology, 52*, 932–941.

Wilming, N., Harst, S., Schmidt, N., & König, P. (2013). Saccadic momentum and facilitation of return saccades contribute to an optimal foraging strategy. *PLoS Computational Biology, 9*, e1002871.

Wu, E. X. W., Gilani, S. O., van Boxtel, J. J. A., Amihai, I., Chua, F. K., & Yen, S.-C. (2013). Parallel programming of saccades during natural scene viewing: Evidence from eye movement positions. *Journal of Vision, 13*(12):17, 1–14, http://www.journalofvision.org/content/13/12/17, doi:10.1167/13.12.17. [PubMed] [Article]